Process Comparison for Subpopulations of Patients with Heart Failure

Eliza Stel
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
e.w.stel@student.utwente.nl

ABSTRACT

Different patients might need different treatments, even if they are admitted for the same condition. The current models for treating patients do not sufficiently take into account how treatments differ per patient group, even though this is useful information for care givers. To find some of the differences in treatments, this research uses process mining to visualize the differences in procedures for patients with heart failure. Process mining uses data from an information system to make process models which show the activities taken by the people within the information system. For this research the information system is the MIMIC-III database, which holds patient data from the intensive care. Process models of patients of different age groups and blood pressure measurements are made. These process models are compared against each other using BPMNDiffViz and human judgement. From this I found that the oldest age group of patients above 65 years have the most differing treatment compared to other age groups. Secondly, the patient group with the highest blood pressure reading upon admittance (blood pressure above 160/90 mm Hg) differed the most from all other groups, as it had the least amount of procedures. These results indicate that there are distinct differences between treatments of different patient groups, future research could show how treatments differ for other patient groups, or go more in-depth and work with a hospital to implement possible changes based on the results.

Keywords

process mining, process comparison, subpopulation, heart failure, ICU, healthcare

1. INTRODUCTION

A lot of research has been done on risk factors for heart failure (HF) and how they affect the chance of getting a heart disease. In precision medicine disease prevention and treatment is personalized to the patient based on their genes, environment and lifestyle. For HF research is being done into the different omics (genomics, pharmacogenomics, epigenomics, proteomics, metabolomics, and microbiomics) and how they relate to the diagnosis, development, progression and treatment of HF. This research has a lot of promising results, but as it is a very large and complicated subject there is still a lot to be done within precision medicine [3]. Thus there is still a need for showing how care paths for patients with different risk factors differ and can be improved upon, where the care path defines which different procedures and treatments are best practice for a specific patient group.

Medical experts are interested in knowing the differences in care paths followed by different patient groups, as this can show how patient get different treatments and if some treatment needs to be adjusted [13]. Furthermore, by knowing the differences between the groups it is possible to optimize the care processes for each subpopulation. Herein, subpopulations are groups of patients who share a common trait, such as all patient older that 80. To find the differences in care paths, process mining can be used. Process mining models the events taken within an organization, in this case the Intensive Care Unit (ICU) of hospitals, and thus gives insight into the reality of the processes, which can be very different from expected processes and standard models. With these insights processes can be optimized while reducing the costs of these processes [12].

In this research there is a focus on patients with HF admitted to the ICU as HF is in the top 5 reasons why a patient is admitted to the hospital [6]. On top of that, one in five patients with HF is admitted to the ICU. This results in a large amount of patients admitted to the ICU, which is a very high-cost setting [20]. Also, as HF prevalence increases with age, while the life expectancy grows this means more people with lasting effects of HF need more hospital care and possible re-admissions to the hospital. Thus there is a need to improve the care systems to also create better care systems which can help more people [5]. HF is the only disease studied within this research, as it already has many subpopulations and possible differences in care, and including more diseases would have given very little time to analyze the differences in care paths.

The aim of this research is to discover the process models of patients with heart failure from 2 different subpopulation categories, these categories based on age and blood pressure. These process models will be compared to show the *differences between treatment procedures given to patients with HF from different subpopulations*. Based on this aim the research question is formulated as follows: What are the differences between treatment processes in the ICU followed by patients with HF from different subpopulations?

To answer the research question, two subquestions have been defined which will be answered first.

1. What are suitable subpopulations for patients with

35th Twente Student Conference on IT July 2nd, 2021, Enschede, The Netherlands.

HF, based on risk factors and chance of readmission?

2. What are the treatment processes for these subpopulations in the ICU?

This paper is structured as follows. The next section will further explain process mining and discuss the tools and data used in this research. Section 3 shows existing related research. Section 4 goes into the methodology used to get the subpopulations and the process models. Section 5 lists the findings from the process model comparison. In the discussion, section 6, the findings are interpreted. Lastly, section 7 is the conclusion and discusses some future work.

2. BACKGROUND

2.1 Process mining

In process mining the goal is to discover, analyze and improve upon real world processes. Data from existing systems can be used to model the processes within an organization, thus showing the actual behaviour of people and the organization which can be compared against the expected behaviour. This can lead to more insights and help improve the processes within the organization, making them more efficient [12].

Process mining has already been successfully applied in different domains such as banking, logistics and smart diagnostics. Unlike with business processes, in healthcare processes it is not possible to apply rigorous business process and workflow management techniques, as healthcare processes need to be more flexible and have many exceptions based on the needs of the patients. In healthcare process mining can be used to bring many improvements upon the organization [12]. Process models can be used to make sure that procedures are firmly understood by care professionals and stakeholders, and make processes more efficient and thus cut costs. Furthermore a better understanding of resources and patients can be developed, insights into collaboration with peers can be given, identifying where long waiting times and bottle necks are, and being able to predict a patient's necessary care based on previous cases [19].

2.2 Tools

For process mining the open-source tool ProM 6.10 was used, which is a program developed for process mining and can be extended with the use of plug-ins. In healthcare ProM is the most used process mining tool [19]. Version 6.10 of ProM is specifically meant for researchers, as this version will not be updated and thus the research results can be replicated [17].

The following plug-ins for ProM were used: convert CSV to XES, Inductive Visual Miner and Convert Petri net to BPMN diagram

The tool BPMNDiffViz [9] was used to help visualizing differences between process models. BPMNDiffViz is a web-based tool developed to find and visualize differences between business processes.

2.3 Data

The MIMIC-III v1.4 database has been used in this research. This database is freely-available and consists of deidentified patient data from the ICU at Beth Israel Deaconess Medical Center [10]. To gain access to this data the required online course on data use and protection was completed and the data use agreement was signed.

The MIMIC database contains data from over 46.000 patients admitted to the ICU from 2001 til 2012. This makes the database well-suited for research, as it contains a large and diverse population of patients [10].

3. RELATED WORK

The paper *Process mining in healthcare: A literature review* [19] discusses the usefulness of process mining in healthcare, and gives an overview of for which ends process mining in healthcare can be used. Furthermore, this paper gives a summary of which methods and tools were used in process mining in healthcare.

The book *Process Mining in Healthcare - Evaluating and Exploiting Operational Healthcare Processes* [12] covers the basis of numerous question which arise when wanting to use process mining in healthcare, specifically focusing on organizational healthcare processes. The book showcases how process mining can be used in healthcare and what problems one can run into while process mining.

Automatic Process Comparison for Subpopulations: Application in Cancer Care [14] applies process mining in the comparison of the treatments of patients from different subpopulations. This paper focuses on cancer patients and uses datasets of MIMIC-III and the Hospital Group Twente. In this paper the process models of different subpopulations are compared using graph comparison methods, including a graph edit distance and cross-log conformance, furthermore a user study was performed to check the results of the graph comparison methods.

BPMNDiffViz is also used in *BPFlexTemplate: a software tool to derive flexible process model templates* [8]. In this paper BPMNDiffViz is extended to also derive flexible template models from two process models. These models can be stored and used for further process comparisons.

As this research is written there is no research yet done on using process mining for showing the differences between subpopulations of patients with HF.

4. METHODOLOGY

The methodology of this research is based on the proposed methodology in [15]. This methodology, called Process Mining Project Methodology in Healthcare (PM^{2}HC), has been used as it is developed for process mining in healthcare, which fits with this research. The paper defines 6 phases: planning, extraction, data processing, mining and analysis, evaluation and lastly improvement and support.

During the planning phase the goals of the research are defined and managing expectations of the stakeholders involved in the research. During this phase the research questions also need to be determined.

The extraction phase deals with extracting the necessary event data from the system. Herein the stakeholders are involved in determining the scope of the research and visual representations are needed to make clear how the information is ordered. The recommended place for data extraction is the billing system.

During the data processing phase the logs which will be used for process mining will be made. The recommended elements in a log are: patient's ID, activity name, department and timestamp.

During mining and analysis the process models are generated and the goal is to gain insight into the processes. Model statistics can be used to compare differences between model or between expected behaviour and the executed behaviour.

With the evaluation phase the goal is to define improvement ideas based on the findings from the analysis. Herein the stakeholders once again play a role, in this case to judge if the conclusions and improvement ideas align with the project's objective. During this phase the results also need to be interpreted to understand the process models and results from the analysis.

In the last phase the stakeholders decide the desired path to implement the improvement ideas. As this research was performed without collaboration with a hospital, this phase will not be executed.

4.1 Phase 1: Planning

The planning phase consisted of making the research proposal. During this phase the course to become a credentialed user of PhysioNet was also completed and the application to get access to the MIMIC-III database was turned in.

4.2 Phase 1.5: Deciding on subpopulations

As finding which subpopulations are suitable for HF will not be answered by means of process mining, it does not fit into the PM^{2}HC. Instead, to find possible subpopulations of patients with HF, literature research was done. Here was focused on finding what groups of people are at risk of getting HF and groups which have a higher chance of being readmitted to the hospital. This methodology has been chosen as the aim of this research is to find the differences in care paths of patients with different risk factors.

The report of heart disease and stroke statistics [23] holds a chapter about the statistics of HF, including the prevalence, incidence, lifetime risk, risk factors and mortality, among other statistics. In this chapter there are many factors discussed which influence the risk of getting heart failure, such as high blood pressure, diabetes, gender, age and ethnicity.

The second paper used focused on subpopulations of patients with HF who were readmitted to a hospital at a later date, also including multiple re-hospitalizations for non-HF admissions to include the significance of a comorbidity burden. If a patient has a comorbidity burden then that means they have multiple diagnoses, such as HF and chronic kidney disease. The subpopulations that were found in this research where, among others, elderly patients with a higher comorbidity burden, patients admitted to the ICU and gender [5].

Based on the different subpopulations in the research, two subpopulations have been chosen to focus on in this research to keep the research within the allotted time frame. These subpopulations are age and blood pressure (BP). Age has been chosen because they prevalence of HF increases drastically the older the population group is. The prevalence of HF ranges from around 1 percent for people under 45, around 6 percent for people between 45 and 65, and around 11 percent for people above 65 [7]. For the BP the risk of getting HF increases with a high BP, where the risk of getting HF is 1.6 times higher when a person has a BP over 160/90 mm Hg compared to a person with a BP below 120/90 mm Hg [23]. This leads to a total of 6 subpopulations: age below 45, between 45 and 65 and above 65, and BP below 120/90 mm Hg, between 120/90 and 160/90 mm Hg and above 160/90 mm Hg.

4.3 Phase 2: Extraction

The data was downloaded from Physionet once the application to access the MIMIC-III database was approved. The data [16] was loaded into a local database so that the PostgreSQL queries could be executed in PgAdmin 4. For visual representation of the data a schema generated SchemaSpy was used [21].

In accordance with the advice from PM^{2}HC of using the billing system for the data extraction [15], the table that was used for the procedures was the 'CPTEVENTS' table, which holds the procedures based on the current procedural terminology (CPT) codes to facilitate billing for the procedures performed.

4.4 Phase 3: Data processing

To get the process models first the tables had to be made of the specific subpopulations, this was done pgAdmin 4.

The MIMIC-III database work with ICD-9 codes for diagnoses, in ICD-9 all codes for heart failure start with '428'. So for this research the data was filtered to only use admissions where the patient was diagnosed with at least one form of heart failure, to ensure all forms of heart failure are included.

To make the tables of all subpopulations, PostgreSQL queries were used. The patients age was gotten through their date of birth and date of admission, as no age was registered in the database. The BP categories were gotten through the first BP measurement upon admittance, as the aim of this research was to see the differences in care paths for subpopulations where the hospital would be able to determine which subpopulation a patient belonged to based on measurements and available information. It would also be interesting to see how an extended period of high blood pressure would affect the care path for future research, but as it is not always possible to know how long a patient has had high blood pressure upon admittance, this was not included in this research.

To make process models the following things need to be defined: cases, events, start time and completion time. One case was for this research one admission, and the events were the procedures based on the CPT names. As the procedures did not have a timestamp, the 'ticket_id_seq' was used, which gave the sequence all procedures were executed in.

Except for ventilation, all procedures only had a sequence number which showed in what order the procedures were done, and not a timestamp. Ventilation only had a timestamp, and because of this all procedures which consisted of ventilation were filtered out, as it was impossible to match the ventilation up with the other procedures.

4.5 Phase 4: Mining and Analysis

The tables of the different subpopulations with the necessary patient data and procedures were put into ProM, where the tables first had to be converted from a CSV file to a XES file using an extension of ProM. This was done as PgAdmin exports CSV files and all ProM mining tools works on XES files.

The CPT names consist of two levels of abstraction. Here the top-level was used, which defines 8 different kinds of procedures. The second level of abstraction gave a more detailed description of the procedure, but this level consisted of 143 events, which was too large to compare in this research.

A few procedures did not have a name in the database. These procedures are kept in the used dataset because in all the process models containing these procedures, they showed up consistently in the same place. Because these unnamed procedures are consistent among the models it is less likely that these events are a simple error, as with an error it would be more likely that it occurs at different times and not consistently after the same procedures. The procedures are shown in the process model with the name 'NULL'.

To make the process models the Inductive Visual Miner was used, which is a robust and user friendly miner and is capable of dealing with noise. It also showed how many patients went into a procedure and how many times that procedure was accessed in total [11]. This information is very useful as it helps to see if a any changes are also significant.

In Inductive Visual Miner the paths slider was set to 90%, to filter out small deviations in the paths. As there are many exceptional behaviours in healthcare and flexibility is necessary [12], these small deviations do not help with seeing the overall differences between subpopulations. The activities slider was put to 1, to make sure that all procedures would be included in the model.

Process models for subpopulations 'BP above 160/90 mm Hg' and 'between 45 and 65 years' are added in appendix A. Due to the size of the models and the maximum page count requirement of the 35th Twente Student Conference on IT the other models are not included.

Furthermore, in ProM the dotted chart was also gotten of all subpopulations. This gave some more statistics, such as the average amount of procedures necessary per patient, and the standard deviation of this average number. These statistics are noted in table 1.

In preparation of comparing the process models they were also made in the Business Process Model Notation (BPMN) 2.0 format. To get the models in the right format the process models were exported as a Petri net and converted using the plug-in 'Convert Petri net to BPMN diagram'. This was done to ensure the BPMN models matched up with the inductive visual miner models and did not differ due to using a different miner to get the Petri net or BPMN models.

4.6 Phase 5: Evaluation

To compare the different process models BPMNDiffViz [9] was used, as it shows clearly the differences between models and can calculate the graph edit distance (GED). The GED defines the number of edits needed to transform one graph into another, where every type of edit can have a different weight[4].

Although BPMNDiffViz only accept models in the BPMN 2.0, which was developed specifically for business processes, they also work for care processes [1].

BPMNDiffViz has implemented 6 different GED algorithms. To decide which algorithm to and weights to use for the edits, the paper by Skobtsov and Kalenkova [22], where they test 5 different GED algorithms on performance and precision, is used. In this paper it is stated that the tabu algorithm has the best precision, with all the edit weights are set to 1 among the different algorithms. Thus, in this research the tabu search was used with the same settings, namely the weight for all edits set to 1 and the maximum number of expansions and size of the tabu list set to 100.

The only algorithm in BPMNDiffViz that was not tested by Skobtsov and Kalenkova was the genetic algorithm. According to [18] this algorithm is very fast and thus suitable for large models, but it only approximates the GED. As the models for this research only have a maximum of 8 procedures, this algorithm was not necessary.

To get the GED the models were uploaded to BPMNDiffViz and with the tabu algorithm the models were compared in sets of two. The comparisons were done among a single risk factor, so the different age groups were compared against each other and the BP groups were compared against each other. This resulted in a process model marked with red for elements deleted and green for elements added in BPMNDiffViz. In figure 1 an example of the visualization is shown of the model 'BP between 120/90 and 160/90 mm Hg' which was compared to 'BP below 120/90 mm Hg'. As 'BP between 120/90 and 160/90 mm Hg' has the most elements out of the two, the elements which are not in 'BP below 120/90 mm Hg' are marked in green. With these markings the procedures were compared and noted down. All results gathered from these comparisons are written out in section 5.

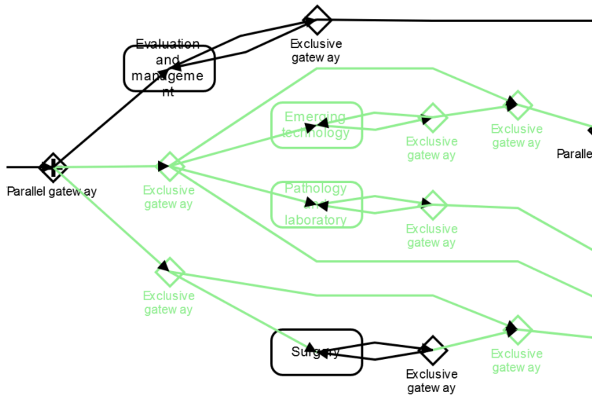

Figure 1. Part of visualization of comparison 'BP between 120/90 and 160/90 mm Hg' and 'BP below 120/90 mm Hg'

All the procedures were compared to where they were in the models. Here the following things were checked: if procedures are parallel, sequential or exclusive. Parallel procedures meant that patients had access to all procedures that are parallel to each other, and it did not matter in which order they were accesses. For instance, if the procedures 'medicine' and 'surgery' are on parallel paths, then patients could either receive 'medicine' first or 'surgery' first. Exclusive meant that patient had only access to one of the procedures that was exclusive, for instance, if 'emerging technology' is exclusive of 'pathology and lab', patients either have 'emerging technology' or 'pathology and lab', not both procedures. If procedures were sequential, then the procedures were accessed in the order defined in the model. For instance, if 'radiology' is sequentially before 'anesthesia' then patients first have to receive 'radiology' before they can receive 'anesthesia'.

5. FINDINGS

Among the process models there were some procedures that were consistent in all models. The first one was 'evaluation and management', which was always in its own *parallel* path and was accessed by all patients, often multiple times accessed. The second similarity was that if a patient has 'anesthesia', they also have 'radiology'. The third was that if there was a 'NULL' procedure, it would be at the end of the model in *parallel* with 'evaluation and management' and *sequentially after* all the other procedures. This is shown in figure 2, which is a detail of the process model from 'above 65 years', where it is visible that the 'NULL' procedure is at the end of the model in parallel with the top line which holds 'evaluation and management' (not shown in the figure). Fourthly, for all subpopulations except 'above 160/90 mm Hg' the procedure 'surgery' was

Table 1. **Statistics subpopulations**

	<45 years	45-65 years	>65 years	BP <120/90	BP 120/90-160/90	BP >160/90
# patients	458	3057	8447	5895	4641	1425
# procedures	6	7	8	8	8	5
Avg # procedures accessed	22.20	19.35	13.06	16.20	14.55	15.27
Std. dev procedures accessed	24.70	24.46	15.77	20.43	18.68	19.27

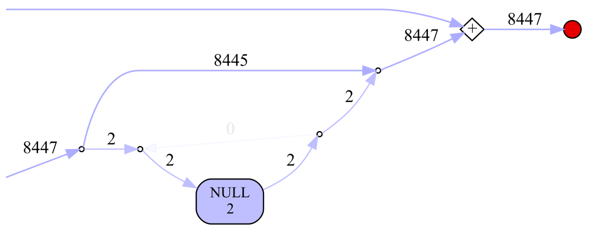

Figure 2. NULL procedure always at the end of the model

sequentially before 'NULL' if applicable and *parallel* to all other procedures.

All the models also contained at least 5 distinct procedures, which were: 'evaluation and management', 'anesthesia', 'radiology', 'surgery' and 'medicine'. Any additional procedures were either 'emerging technology', 'pathology and laboratory' or 'NULL'.

5.1 Subpopulation: Age

5.1.1 Graph edit distances

The GED for the age subpopulations were: 35 edits for 'below 45 years' and 'between 45 and 65 years' (appendix A, figure 5), 48 edits for 'below 45 years' and 'above 65 years' and 49 edits for 'between 45 and 65 years' and 'above 65 years'.

5.1.2 Differences procedures and paths

The procedure **'emerging technology'** was only accessed in the subpopulation 'above 65' by 14 patients, where it was *sequentially before* 'radiology', 'anesthesia' and 'pathology and laboratory'. In the other two age subpopulations it was never accessed.

The procedure **'NULL'** was accesed by subpopulation 'above 65 years' by 2 patients and by subpopulation 'between 45 and 65 years' by 1 patient, and not accessed by the subpopulation 'below 45 years'.

'Pathology and laboratory' was *sequentially before* 'radiology and anesthesia' for subpopulation 'below 45' where 'pathology and laboratory' was accessed by 3 patients and 'radiology' accessed by 19 patients. For subpopulation 'between 45 and 65' 'pathology and laboratory' (4 patients) was *exclusive* of 'radiology' (187 patients) and 'anesthesia'. For subpopulation 'above 65 years' 'pathology and laboratory' (13 patients) was *sequentially after* 'radiology' (531 patients) and 'radiology'.

'Medicine' for subpopulations 'below 45' and 'between 45 and 65' was *sequentially after* 'pathology and laboratory', 'radiology' and 'anesthesia', where 'medicine' is accessed by 125 patients in 'below 45' and 835 patients in 'between 45 and 65'. In 'between 45 and 65' 'medicine' can also be skipped, whereas in 'below 45' that is not possible. For subpopulation 'above 65', 'medicine' (1047 patients) is in *parallel* 'pathology and laboratory', 'radiology' and 'anesthesia'.

For 'below 45' **'radiology'** (19 patients) is *sequentially before* **'anesthesia'** (2 patients). In subpopulations 'between 45 and 65' and 'above 65' 'radiology' is in *parallel* with 'anesthesia', with 'radiology' accessed by 187 patients for 'between 45 and 65' and by 531 patients for 'above 65', and 'anesthesia' accessed by 13 patients for 'between 45 and 65' and by 29 patients for 'above 65'. Figure 3 show the procedures in the models of 'below 45' and 'above 65', here it is visible that 'radiology' and 'anesthesia' are sequential in 'below 45' while in 'above 65' they are parallel.

5.2 Subpopulation: Blood pressure

5.2.1 Graph edit distances

The GED for the BP subpopulations were: 23 edits for 'below 120/90 mm Hg' and 'between 120/90 and 160/90 mm Hg', 54 edits for 'below 120/90 mm Hg' and 'above 160/90 mm Hg' (appendix A, figure 4) and 51 edits for 'between 120/90 and 160/90 mm Hg' and 'above 160/90 mm Hg'.

5.2.2 Differences procedures and paths

Procedure **'NULL'** was accessed by 3 patients for 'below 120/90 mm Hg' and by 1 patient for 'between 120/90 and 160/90 mm Hg', and never for 'above 160/90 mm Hg'.

For 'below 120/90 mm Hg' **'pathology and laboratory'** (7 patients) was *sequentially after* 'radiology' (383 patients) and 'anesthesia'. For 'between 120/90 and 160/90' 'pathology and laboratory' (6 patients) was *exclusive* of 'radiology' (258 patients) and 'anesthesia'. In the process model for 'above 160/90 mm Hg' 'pathology and lab' was never accessed.

Procedure **'emerging technology'** (6 patients) was *exclusive* of 'radiology' (383 patients) and 'anesthesia' in subpopulation 'below 120/90 mm Hg'. For 'between 120/90 and 160/90 mm Hg' 'emerging technology' (7 patients) was *sequentially before* 'radiology' (258 patients) and 'anesthesia'.

'Medicine' (865 patients) in *parallel* with 'surgery', 'radiology', 'anesthesia', 'pathology and laboratory' and 'emerging technology' for subpopulation 'below 120/90 mm Hg'. For 'between 120/90 and 160/90 mm Hg' 'medicine' (610 patients) is *sequentially after* 'radiology', 'anesthesia', 'pathology and laboratory' and 'emerging technology'. For 'above 160/90 mm Hg' 'medicine' (221 patients) is in *parallel* with 'surgery' and *sequentially after* 'anesthesia' and 'radiology'.

For both 'below 120/90 mm Hg' and 'between 120/90 and 160/90 mm Hg' **'radiology'** (383 and 258 patients, respectively) was in *parallel* with **'anesthesia'** (17 and 22 patients, respectively). For subpopulation 'above 160/90 mm Hg' 'radiology' (44 patients) was *sequentially before* 'anesthesia' (2 patients). For 'above 160/90 mm Hg' it was never accessed.

For both 'below 120/90 mm Hg' and 'between 120/90 and 160/90 mm Hg' **'surgery'** (2246 and 1526 patients, respectively) is in *parallel* with 'pathology and laboratory', 'emerging technology', 'radiology', 'anesthesia' and

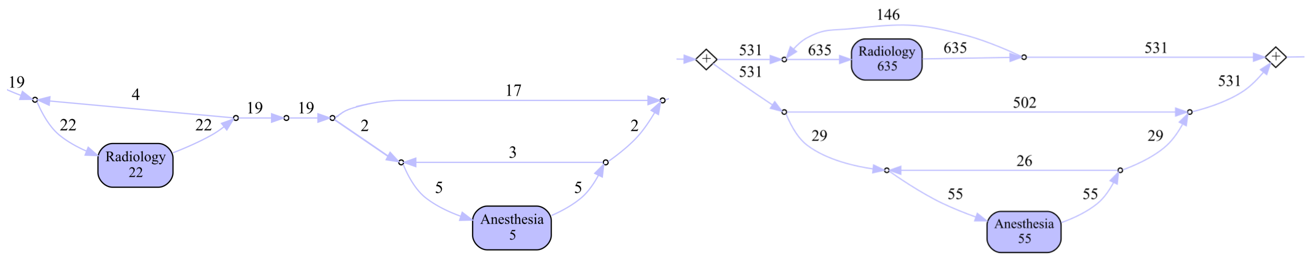

Figure 3. Model comparison of 'radiology' and 'anesthesia' for models 'below 45' (left) and 'above 65' (right)

'medicine'. For 'above 160/90 mm Hg' 'surgery' (827 patients) is in *parallel* with 'medicine' and *sequentially after* 'anesthesia' and 'radiology'.

6. DISCUSSION

Not much is known how different risk factors influence the care a patient receives at a hospital. To give more insight herein this research focuses on showing the differences in care received by patients with heart failure from different age groups and BP measurements. Process models of patients of the subpopulations were compared to see the differences in which procedures patients received and in which order the procedures were given.

The biggest differences found in the care was for the subpopulation with BP above 160/90 mm Hg and the subpopulation of patients older than 65 years.

The subpopulation 'BP above 160/90 mm Hg' differs the most from the other subpopulations, as it only has 5 procedures, compared to the 8 procedures in the other subpopulations defined by BP measurements. Furthermore, all other subpopulations have the procedure 'surgery' in parallel with all other procedures, except for the 'NULL procedure, but the subpopulation 'above 160/90 mm Hg' has surgery sequentially after 'radiology' and 'anesthesia', which means that patients that have 'radiology', 'anesthesia' and 'surgery' first need to go through 'radiology' and 'anesthesia' before getting 'surgery'.

The subpopulation 'above 65 years' differs from the other subpopulations defined by age because it is the only one with the procedure 'emerging technology'. This meant that newer procedures and technologies were only used for older patients. Furthermore, 'medicine' had its own parallel path, meaning that elderly patients needed medical services and procedures earlier in the treatment process than younger patients.

Although the 'NULL' procedure was not in all subpopulations present, I think this might not necessarily be due to the difference in care a subpopulation needs but due to difference in number of patients. The two largest subpopulations, 'above 65 years' (8447 patients) and 'BP below 120/90 mm Hg' (5895 patients), only have 2 and 3 patients with the 'NULL' procedure, whereas the two smallest subpopulations, 'below 45 years' (458 patients) and 'BP above 160/90 mm Hg' (1425 patients), do not have the 'NULL' procedure. As there is a large size difference of these subpopulations it is very much possible that the 'NULL' procedure is not in the smaller subpopulations because these subpopulations are too small.

Among all subpopulations the procedure 'pathology and laboratory' tended to be in a different place in the process models. It was either before, after, or exclusive of 'radiology' and 'anesthesia'. This could possibly mean that only a few patients needed both 'pathology and laboratory' and 'radiology'/'anesthesia' and thus when a patient needs both this is decisive for the order of the procedures. Another possibility is that 'pathology and laboratory' is needed differently depending on the subpopulations. Currently it is not possible to see how many patients had both 'pathology and laboratory' and 'radiology'/'anesthesia', so no conclusive reasoning can be chosen.

From the findings follow that while there are some similarities among all subpopulations, every subpopulation also has its own distinct paths. This is not surprising as in [2] it was shown that elderly patients with heart failure need different care than younger patients, as they react differently to some medicine and procedures and are more likely to have multiple conditions which need to be treated.

6.1 Limitations and future work

This research only used data of one hospital, so it is possible that there is a bias in the data consisting of the practices of this hospital, which could differ from the practices of other hospitals. In the future it could be looked into how practices differ per hospital and gather data of multiple hospitals. Future research which would also implement improvement ideas in a hospital should look at the specific practices of that hospital to ensure the improvements fit with that specific setting.

A limitation of this research was that there were no timestamps available for the procedures, for future research it would be interesting to work with data with timestamps and see where there are bottlenecks and long waiting times.

Future research could expand on this research and include more subpopulation, or go into more detail by using subevents of the procedures. As this research has shown that there are differences in treatments for different patient groups, it would be beneficial to further show the differences and work with a hospital to implement improvement ideas.

7. CONCLUSION

In this research the differences between treatments of patient groups with heart failure were shown. Here different subpopulations were decided on based on how much different risk factors influenced chance of getting heart failure and being readmitted for heart failure. In this research the subpopulations focussed on where age and blood pressure. From these subpopulations were process models made, and I compared those using BPMNDiffViz and checking where the procedures were relative to each other. From the comparison followed that the patients older than 65 had the most different treatment among the age groups, as it was the only age group that had emerging technology as a procedure and medicine was in parallel with all other procedures, instead of after radiology and anesthesia. The subpopulation of patients with a blood pressure above 160/90 mm Hg differed the most from all other subpopulations, as it had the smallest amount of procedures and did not have surgery in parallel with all other procedures.

A major limitation of this research was that there were no timestamps available of the procedures, in future research a data set needs to be found with timestamps to see where long waiting times are and bottlenecks. Furthermore the MIMIC-III dataset is very large and complex, and although I tried to ensure no cases were missed, it is possible that some cases got lost during processing.

As there are clear differences between treatments of different subpopulations, future research can go more in-depth and work on improvement ideas to apply in a hospital and also work with different subpopulations.

8. ACKNOWLEDGMENTS

I am thankful for the help of my supervisor, Faiza Bukhsh, and Syeda Sohail for giving valuable input and feedback into this research. Without them this research would not have come as far as it has.

9. REFERENCES

[1] J. Chae, B. H. Park, M. Jones, M. Ward, and J. Nebeker. Converting clinical pathways to bpm+ standards: A case study in stable ischemic heart disease. In *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, volume 2020-July, pages 453–456, 2020.

[2] M. Colvin, N. K. Sweitzer, N. M. Albert, R. Krishnamani, M. W. Rich, W. G. Stough, M. N. Walsh, C. A. Westlake Canary, L. A. Allen, M. R. Bonnell, P. E. Carson, M. C. Chan, M. G. Dickinson, D. L. Dries, G. A. Ewald, J. C. Fang, A. F. Hernandez, R. E. Hershberger, S. D. Katz, S. Moore, J. E. Rodgers, J. G. Rogers, A. R. Vest, D. J. Whellan, and M. M. Givertz. Heart failure in non-caucasians, women, and older adults: A white paper on special populations from the heart failure society of america guideline committee. *Journal of cardiac failure*, 21(8):674–693, 2015.

[3] S. Cresci, N. L. Pereira, F. Ahmad, M. Byku, L. De Las Fuentes, D. E. Lanfear, C. M. Reilly, A. T. Owens, and M. J. Wolf. Heart failure in the era of precision medicine: A scientific statement from the american heart association. *Circulation: Genomic and Precision Medicine*, 12(10):458–485, 2019.

[4] R. Dijkman, M. Dumas, B. Van Dongen, R. Krik, and J. Mendling. Similarity of business process models: Metrics and evaluation. *Information Systems*, 36(2):498–516, 2011.

[5] F. Gasperoni, F. Ieva, G. Barbati, A. Scagnetto, A. Iorio, G. Sinagra, and A. Di Lenarda. Multi-state modelling of heart failure care path: A population-based investigation from italy. *PLoS ONE*, 12(6), 2017.

[6] HCUP Fast Stats. Healthcare Cost and Utilization Project (HCUP). April 2021. Agency for Healthcare Research and Quality, Rockville, MD. www.hcup-us.ahrq.gov/faststats/national/inpatientcommon-diagnoses.jsp.

[7] P. A. Heidenreich, N. M. Albert, L. A. Allen, D. A. Bluemke, J. Butler, G. C. Fonarow, J. S. Ikonomidis, O. Khavjou, M. A. Konstam, T. M. Maddox, G. Nichol, M. Pham, I. L. Piña, and J. G. Trogdon. Forecasting the impact of heart failure in the united states a policy statement from the american heart association. *Circulation: Heart Failure*, 6(3):606–619, 2013.

[8] L. Ilahi, R. Martinho, and S. A. Ghannouchi. Bpflextemplate: A software tool to derive flexible process model templates. In *Procedia Computer Science*, volume 121, pages 1096–1103, 2017.

[9] S. Y. Ivanov, A. A. Kalenkova, and W. M. P. Van Der Aalst. Bpmndiffviz: A tool for bpmn models comparison? In *CEUR Workshop Proceedings*, volume 1418, pages 35–39, 2015.

[10] A. E. Johnson, T. J. Pollard, L. Shen, L. H. Lehman, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark. Mimic-iii, a freely accessible critical care database. *Scientific data*, 3:160035, 2016.

[11] S. Leemans. Robust process mining with guarantees. May 2017. Proefschrift.

[12] R. S. Mans, W. M. van der Aalst, and R. J. Vanwersch. *Process Mining in Healthcare - Evaluating and Exploiting Operational Healthcare Processes*. Springer, 2015.

[13] R. S. Mans, W. M. P. Van Der Aalst, R. J. B. Vanwersch, and A. J. Moleman. *Process mining in healthcare: Data challenges when answering frequently posed questions*, volume 7738 LNAI of *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2013.

[14] F. Marazza, F. A. Bukhsh, J. Geerdink, O. Vijlbrief, S. Pathak, M. van Keulen, and C. Seifert. Automatic process comparison for subpopulations: Application in cancer care. *International Journal of Environmental Research and Public Health*, 17(16):1–23, June 2020.

[15] G. B. Pereira, E. A. P. Santos, and M. M. C. Maceno. Process mining project methodology in healthcare: a case study in a tertiary hospital. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 9(1), 2020.

[16] T. J. Pollard and A. E. Johnson. The mimic-iii clinical database. http://dx.doi.org/10.13026/C2XW26, 2016.

[17] Prom 6.10. http://promtools.org/doku.php?id=prom610 (accessed: 01/06/2021).

[18] K. Riesen, A. Fischer, and H. Bunke. *Improving approximate graph edit distance using genetic algorithms*, volume 8621 LNCS of *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2014.

[19] E. Rojas, J. Munoz-Gama, M. Sepúlveda, and D. Capurro. Process mining in healthcare: A literature review. *Journal of Biomedical Informatics*, 61:224–236, 2016.

[20] K. C. Safavi, K. Dharmarajan, N. Kim, K. M. Strait, S. X. Li, S. I. Chen, T. Lagu, and H. M. Krumholz. Variation exists in rates of admission to intensive care units for heart failure patients across hospitals in the united states. *Circulation*, 127(8):923–929, February 2013.

[21] Schemaspy analysis of mimic.mimiciii - all relationships. https://mit-lcp.github.io/mimic-schema-spy/relationships.html (accessed: 12/05/2021).

[22] A. Skobtsov and A. Kalenkova. Efficient algorithms for finding differences between process models. In

Proceedings - 2019 Ivannikov Ispras Open Conference, ISPRAS 2019, pages 60–66, 2019.

[23] S. S. Virani, A. Alonso, H. J. Aparicio, E. J. Benjamin, M. S. Bittencourt, C. W. Callaway, A. P. Carson, A. M. Chamberlain, S. Cheng, F. N. Delling, M. S. Elkind, K. R. Evenson, J. F. Ferguson, D. K. Gupta, S. S. Khan, B. M. Kissela, K. L. Knutson, C. D. Lee, T. T. Lewis, J. Liu, M. S. Loop, P. L. Lutsey, J. Ma, J. Mackey, S. S. Martin, D. B. Matchar, M. E. Mussolino, S. D. Navaneethan, A. M. Perak, G. A. Roth, Z. Samad, G. M. Satou, E. B. Schroeder, S. H. Shah, C. M. Shay, A. Stokes, L. B. VanWagner, N.-Y. Wang, and C. W. Tsao. Heart disease and stroke statistics-2021 update. *Circulation*, 143(8):e254–e743, 2021.

APPENDIX

A. PROCESS MODELS

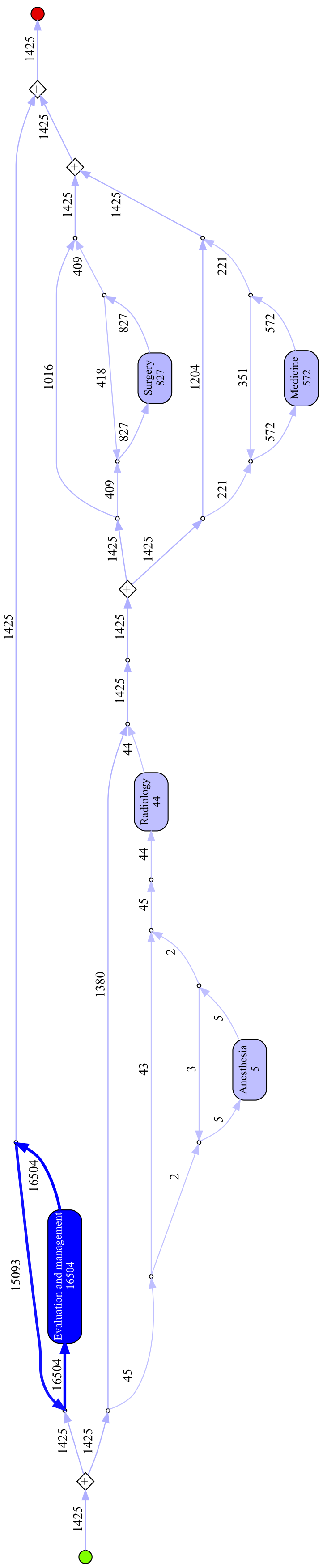

Figure 4. Blood pressure: above 160/90 mm Hg

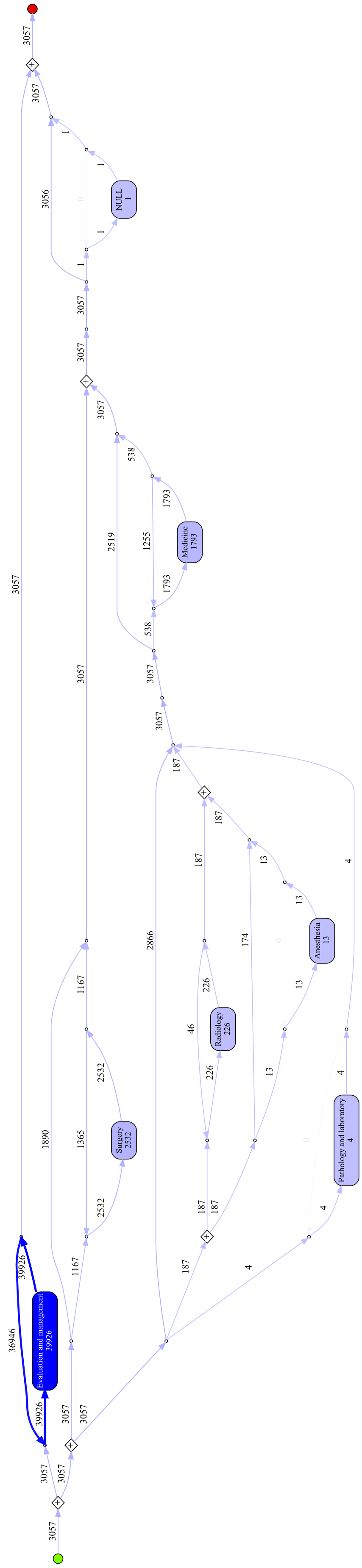

Figure 5. Age: between 45 and 65 years